Performance Evaluation of Sensors on Mobile Vehicles Using a Large Data Repository and Ground Truth

Tsai Hong, Tommy Chang, Ayako Takeuchi, Gerry Cheok, Harry Scott, Michael Shneier
National Institute of Standards and Technology
100 Bureau Drive, Stop 8230
Gaithersburg, MD 20899

ABSTRACT

Progress in algorithm development and transfer of results to practical applications such as military robotics requires standard qualitative and quantitative measurements for performance evaluation and validation. Although the evaluation and validation of algorithms have been discussed for well over a decade, the research community still faces a lack of well-defined and standardized methodology. In this research, we describe three methods for creating ground truth databases and benchmarks using multiple sensors, for use in mobile robotics. The databases and benchmarks provide researchers with high quality data from suites of sensors operating in complex environments representing real-world problems. At NIST, we have equipped a High Mobility Multi-purpose Wheeled Vehicle (HMMWV) with a suite of sensors including a Riegl ladar, GDRS ladar, stereo CCD, several color cameras, Global Positioning System (GPS), Inertial Navigation System (INS), pan/tilt encoders, and odometry[†]. All sensors are calibrated and registered with each other in space and time. This allows a database of features and terrain elevation to be built. Ground truth information is collected through aerial surveys, from maps, by human annotation, and by previous traverses of the terrain by the vehicle. Ground truth may include terrain elevation information, feature information (roads, road signs, trees, ponds, fences, etc.) and constraint information (e.g., one-way streets). We have implemented our a priori database using One Semi-Automated Forces (OneSAF), a military simulation environment. Using the Inertial Navigation System and Global Positioning System (GPS) on the HMMWV to provide indexing into the database, we extract all the elevation and feature information for a region surrounding the vehicle as it moves about the NIST campus. Ground truth for each sensor can be obtained by projecting this information into the sensors' coordinate systems. The main goal of this research is to provide ground truth databases for researchers and engineers to evaluate algorithms for effectiveness, efficiency, reliability, and robustness, thus advancing the development of algorithms.

Keywords: *a priori knowledge, prediction, recognition, processing, data collection, registration, calibration, performance evaluation, ground truth, mobile robots.*

1 INTRODUCTION

Historically, performance evaluation has not been commonly practiced in the perception community. Periodically, efforts are made to persuade researchers to provide performance evaluations that can be substantiated, but only a few take up this challenge. As a result, performance evaluation is ad hoc in general, and quite frequently completely absent from research papers. In Europe, a number of formal programs have been developed that address performance evaluation of vision algorithms. Of these, ECVnet, an association of European vision researchers, had a subcommittee on Benchmarking and Performance Measures[1], although it now appears to be defunct. The German Association for Pattern Recognition (DAGM) established a Working Group on "Quality Evaluation of Pattern Recognition Algorithms", but it, too appears inactive[2]. The International Association for Pattern Recognition has a Technical Committee on Benchmarking & Software, which organizes performance competitions comparing algorithms for particular applications, such as fingerprint identification and document analysis[3]. There have also been a number of workshops on performance characterization and benchmarking of vision systems.

A number of publications address the issue of how to evaluate the performance of vision algorithms and provide examples of careful evaluations of particular algorithms or classes of algorithms. Approaches to performance evaluation can be classified into the following general categories, recognizing that more than one approach may be used in an evaluation.

Comparative Here an algorithm may be compared with others that attempt to address the same image-processing task, or its performance may be compared to "ground truth," or perhaps to human performance[4-9].

[†] Certain commercial equipment, instruments, or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily best for the purpose.

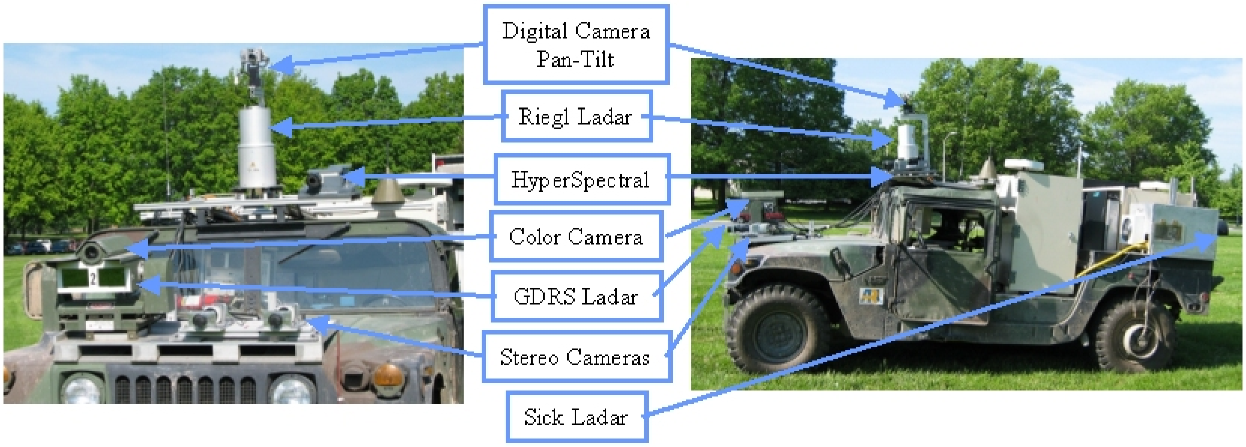


Figure 1 *A view of the NIST HMMWV showing some of the sensors.*

Analytic The theory behind the algorithm is examined to try to determine the limits to its operation. The computational complexity may be derived, or theoretical optimality may be determined under certain constraints. Frequently, the approach makes use of simplified input data to make the analysis feasible[10-13].

Performance The way the algorithm actually performs on test data is measured and execution times with different parameters may be reported[14-16].

Appropriateness to Task The algorithm is shown in the context of a particular application, and the constraints of the task are used to justify the selection of the particular algorithm. The performance of the task as a whole is taken as the evaluation of the algorithm[17;18].

Other, more informal measures include generality and acceptance. Perhaps the only real performance evaluation measure in common use is longevity. Algorithms that are accepted widely and implemented by many people for different applications can be considered good performers.

A large number of papers report excellent performance of their algorithms, based on small data sets. The success of the FERET program[9] has inspired us to take up the challenge of producing a large database of ground truth for the domain of mobile robotics. In this domain, sensors are mounted on the moving vehicle, and the algorithms are constrained to run in real time (i.e., fast enough to provide data to properly control the vehicle). The ground truth that we provide is much more extensive than is typically available, and where human interpretations provide the ground truth, they cover a large number of image sequences because the annotation of the images is performed with computer assistance.

We have developed reliable methods of producing three different kinds of large databases of sensor data with ground truth. One method involves collecting ground truth data using a highly accurate ladar sensor mounted on our instrumented HMMWV. The ladar can characterize large areas of terrain and is registered with cameras that provide color information for each ladar point. The position and time at which each sample is collected is recorded with an Inertial Navigation System (INS) and Global Positioning System (GPS) accurate to a few centimeters. Another set of data was obtained through a high-resolution aerial survey of the grounds of the National Institute of Standards and Technology (NIST) and surrounding area. The survey includes annotations providing labels for all the features. Lastly, we have developed an interactive method of hand-labeling features in image sequences to efficiently generate a large database of ground truth data.

The data sets are used to evaluate performance of algorithms objectively by comparing the output of the algorithms to the expected result derived from the ground truth. Given a large number of ground truth data sets from different environments, statistical evaluations are possible as well as the robust assessment of performance of algorithms.

The main goal of this work is to make our test data and ground truth available for general use, with the hope that it will lead to rapid and significant development of perception algorithms for autonomous mobility. In order to validate the approach we use the data sets to evaluate our own algorithm development.

The NIST HMMWV is a military vehicle modified for the purposes of research and development in mobile robotics. Mounted on the vehicle are racks to hold computers and related equipment, a power generator, and numerous sensors (Figure 1). The sensor mounts are flexible, so that new sensors can easily be added. Sensors include a General Dynamics Robotics Systems (GDRS) imaging ladar mounted on a tilt platform, a color camera mounted on top of the ladar on the tilt platform, and a highly accurate position and orientation system[19].

Other sensors that are commonly used on the vehicle include a pair of color stereo cameras, a Sick line-scan ladar, currently mounted on the back of the vehicle, and a Riegl high-resolution scanning ladar.

2. SENSOR CALIBRATION AND REGISTRATION

Our goal is to provide sets of data that are registered in both space and time. Registration allows information from multiple sensors to be integrated and enables the spatial relationships between successive samples to be computed. Preparing for data collection includes calibrating the sensors and accurately measuring their positions and orientations on the vehicle. Then data are collected from calibrated courses containing known objects to enable the capabilities of each sensor to be quantified.

The cameras are calibrated using Bouguet's method[20]. The ladars are each calibrated using special-purpose methods. For example, the GDRS ladar is calibrated by mounting it on a highly accurate pan-tilt platform. The pointing direction of each laser pixel is determined by moving the ladar until the laser beam for that pixel is centered on a calibration target. The angle at which the laser is pointed for each pixel can then be determined from the pan and tilt position of the platform.

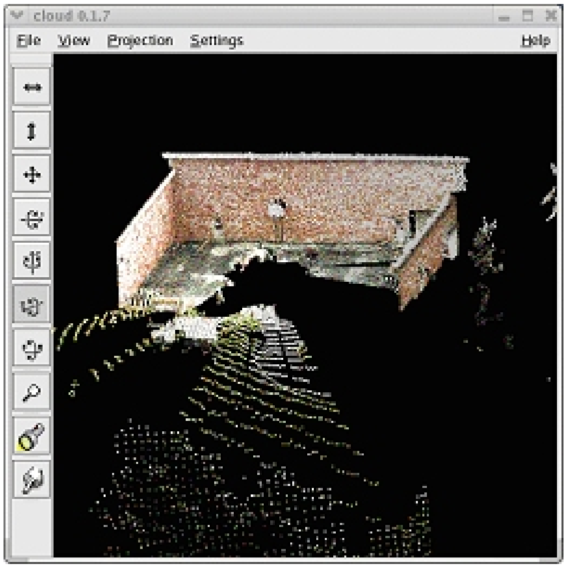


Figure 2 *Data from a scan of orthogonal walls using the Riegl ladar.*

The positions and orientations of the sensors relative to the vehicle coordinate system and to each other are determined using an external measurement system. We use an ArcSecond laser-based site measurement system (SMS) to provide these measurements. For the Riegl ladar, the approach is to park the vehicle in such a way that it faces two orthogonal walls. The Riegl is then used to scan these walls and the ground Figure 2, and a transformation is obtained from the building to the Riegl coordinates. The ArcSecond sensor is then used to determine the HMMWV to ArcSecond transform and the wall to ArcSecond transform. Finally, the Riegl to HMMWV transform can be obtained by matrix multiplication: *Riegl to HMMWV = Riegl to Building * Building to ArcSecond * ArcSecond to HMMWV*. Similar methods are used to locate the other sensors relative to the vehicle.

3. COLLECTING DATA

Data are collected in two primary modes. One is while the vehicle is driving normally, while the other is with the vehicle stationary. Some of the sensors do not run in real time, so can only be used when the vehicle is not moving. The trade-off between the two modes is that while data acquired in real-time approximate more closely the actual driving conditions, they are less accurate and usually of lower resolution than data from the slower sensors used when the vehicle is stopped. The expectation is that this higher resolution data will soon become available in real time as new sensors are developed.

A critical part of data collection for mobile vehicle applications is to record the vehicle position and orientation (pose) and the time at which each data sample is acquired. This enables data collected from multiple sensors to be registered, and also allows the data for a complete mission to be compiled into a reconstruction of all the terrain that was traversed. The vehicle pose and the time are provided by an Applanix navigation unit that combines an inertial component with information from the Global Positioning System (GPS). This unit typically provides real-time data accurate to better than one meter and a few hundredths of a degree. With post-processing, the accuracy is a few centimeters in distance and angular accuracy is a few thousandths of a degree.

To date, data have been collected for two main purposes. The first is to provide a large variety of input data for developing and testing sensory processing algorithms. More recently, a new application has required characterizing terrain and developing measures of difficulty of traversal for robotic vehicles. This has led to a more structured way of collecting data, and is the main reason for needing highly accurate pose and time information and for using high-resolution sensors such as the Riegl ladar.

Collecting data for algorithm development involves driving the vehicle in the way it would normally operate, over terrain representative of the environment in which the vehicle normally operates. The data are acquired as follows. First, the sensors are calibrated and registered, and the navigation system is initialized. Next, a human driver drives the vehicle over the terrain of interest, and the sensors collect all the real-time sensory data simultaneously. The sensors are started simultaneously, and navigation and timing data are collected with the sensory data. Typically, the collection is divided into segments about three minutes in length, mainly for convenience in later processing and storage. Longer periods can also be collected, limited only by the available storage space, which is large (several hundred Gigabytes).

When data are collected for terrain characterization, the process is more methodical. Data have usually been collected on courses laid out for evaluating the capabilities of an experimental unmanned vehicle (XUV) developed under the Army's Demo III program. These courses are defined by a set of GPS waypoints, through which the XUV is supposed to pass as it carries out its mission. Data are collected both for the entire course and for locations that required an emergency stop for the XUV or where the vehicle displayed "interesting" behavior, such as backing up, suddenly changing direction, or performing an unanticipated intelligent maneuver.

Three sets of data are collected for each course. First, the vehicle is driven over the course collecting data with the real-time sensors. Next, the vehicle is moved to the first waypoint on the course. Starting from this point, and moving a fixed distance between samples, scans are taken of the terrain using the Riegl ladar and a digital camera on a pan-tilt unit that captures approximately the same field of view as the Riegl. The scans are not taken at the highest resolution the ladar sensor can measure, but still provide much more accurate information than the real-time sensors. The navigation data are also stored to provide the position and heading of the data collection vehicle at the time the sample is collected. The entire course is sampled in this way. Finally, a set of high-resolution scans is taken of the difficult or interesting locations on the course.

4. ANNOTATING DATA FOR GROUND TRUTH

We first discuss our method for creating ground truth databases for sequences of color image data. It involves a human user, who annotates the data to supply the ground truth. Manually annotating sensor data with ground truth is costly and time consuming. Instead, we have developed a semi-automatic ground truth application that reduces cost and time by requiring only occasional annotation. The user annotates the first image of a sequence by outlining and naming regions of interest (e.g., highway signs, vehicles). The computer then tracks the annotated regions through successive images, and the user observes

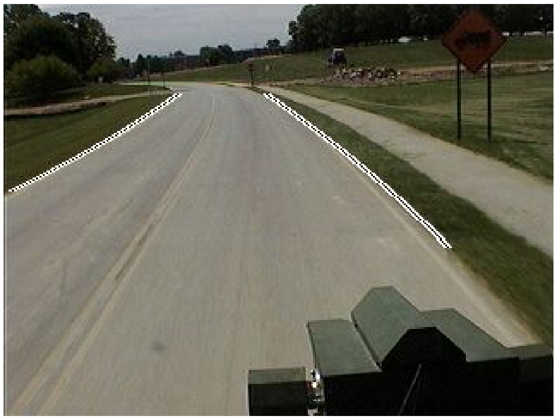


Figure 3 *The first frame of a sequence. The user has drawn the features to be tracked.*

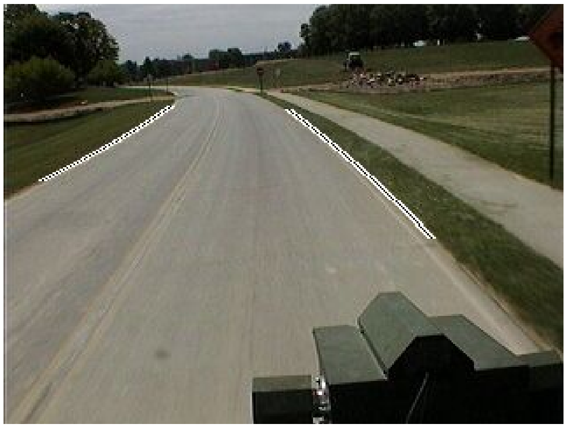


Figure 4 *The computer tracks the features through a sequence of images*

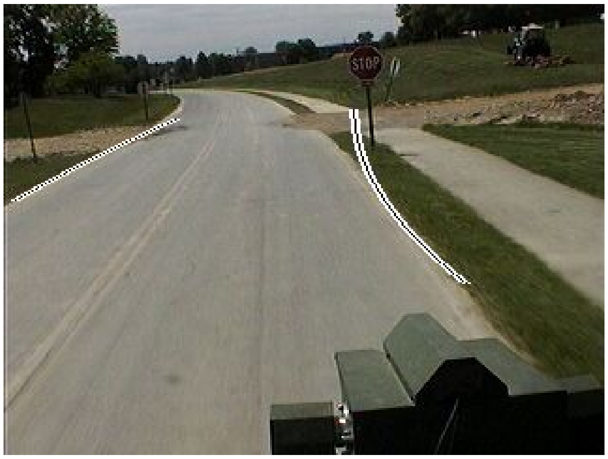


Figure 5 *In this frame, the automatic tracker has drifted enough to require human intervention.*

how well each region is recognized and outlined by the computer. When the annotations start to diverge from the desired regions, the user intervenes and re-identifies the regions, retaining the same names. When new regions appear that the user wants to track, the same process of stopping the computer, annotating the regions, and restarting the tracking is followed. The annotation application can be used to outline regions with curved or polygonal lines, and several tracking algorithms can be used, depending on the objects in the images. The output of this process consists of the names, shapes and position coordinates of the targets in each image

Figure 7 *Example scene from an off-road data set*

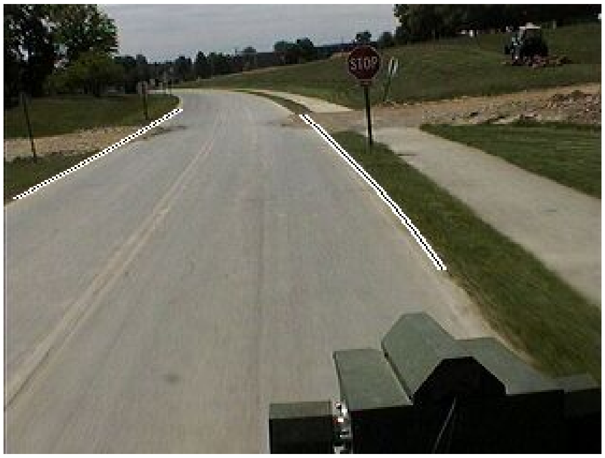


Figure 6 *The user re-initializes the features and automatic tracking continues*

Figure 3 shows the starting frame of a sequence of color images. It shows road edges that were selected by a user constructing the ground truth. Figure 4 shows the results of automatic tracking. The tracking to this point is acceptable, and no user interaction is required. In Figure 5, the tracker is starting to lose the edge of the road. At this point, the user intervenes, selects the road edge again, and the tracking continues (Figure 6)

5. EVALUATING SENSORS

Another approach provides data for evaluating range sensors. It makes use of a high-resolution ladar (Riegl LMS Z210) to construct a map of a region. The map can then be used for evaluating range sensors that have significantly lower resolution than the Riegl. We use a 5 cm x 5 cm spatial resolution grid to construct the ground truth map, but maps can be constructed at different resolutions (finer or coarser). This method has been used to gather ground truth for off-road terrain such as that shown in Figure 7.

Evaluating other range sensors involves mapping their data into the high-resolution grid. The residual of the Riegl data and the other sensor data provides a measure of the performance of the sensor (relative to the Riegl). It is important to note that in order to map data from the sensor under test onto the Riegl data, the positions and orientations of the sensors must be known accurately. The current map resolution of 5 cm x 5 cm corresponds to a spatial tolerance of 5 cm. This method of constructing a map of a region can also measure how much information each successive ladar image adds about the world. The ground truth maps can also be used to evaluate similar maps constructed with stereo algorithms[6]

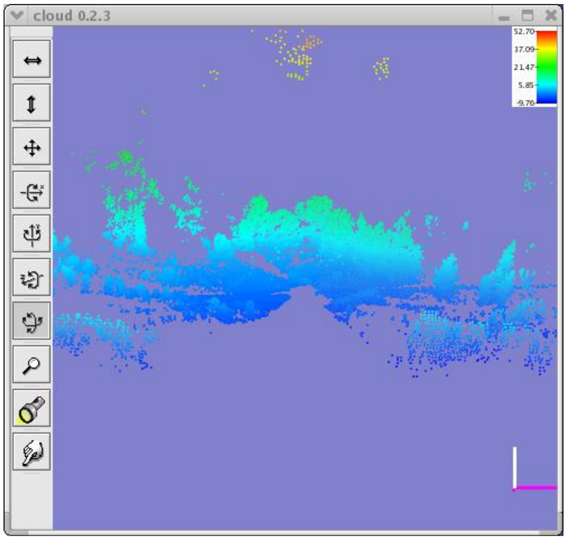


Figure 8 *Range data from the Riegl ladar. Color is used to represent elevation.*

Figure 8 shows the result of scanning a region with the Riegl ladar. Figure 9 shows the sub-region scanned with a different ladar (GDRS). In Figure 10 the two scans are overlaid. The white region shows the mismatch due to the lower resolution and coarser range quantization of the GDRS ladar (with a small component due to registration error).

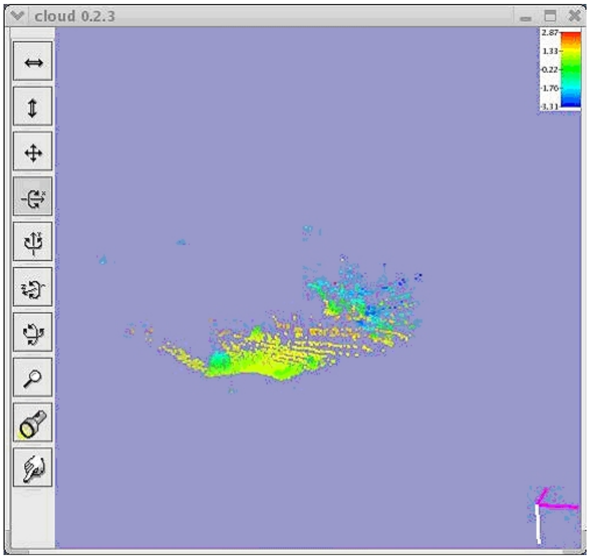


Figure 9 *The sub-region seen by the GDRS ladar, taken from the same position. Elevation is again represented by color.*

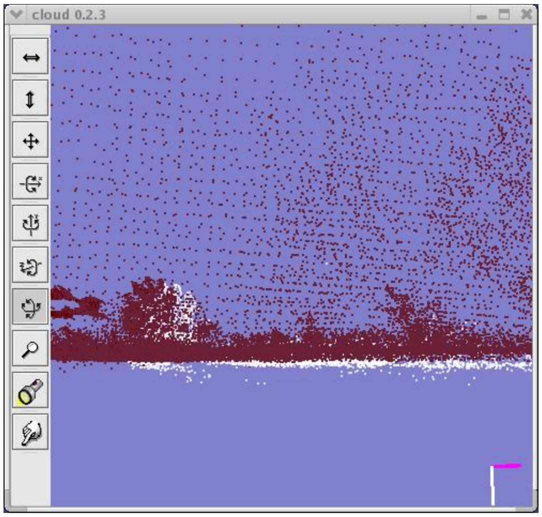


Figure 10 *The result of overlaying the GDRS ladar data on the Riegl data. The difference in measurement of the scene can clearly be seen.*

The third approach to performance evaluation involves constructing a ground truth database of color and range images based on a high-resolution aerial survey combined with data from calibrated ground sensors such as cameras and ladars. In our case, we commissioned a survey of the NIST campus ($234\times10^4\,\mathrm{m}^2$ or 578 acres) and part of the surrounding urban area. The area includes roads, parking lots, traffic signs, buildings, trees, streams, fences, etc., as well as off-road terrain. All of these features are recorded and entered into a database of features and terrain elevation. Ground truth for each sensor can then be extracted from the database based on position and sensor model.

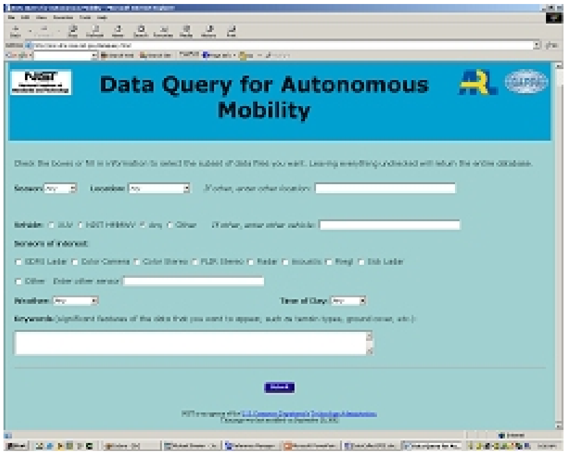
6. DATA STORAGE AND ACCESS

Information about each set of data is stored in a relational database, and the data sets themselves are stored in a large capacity storage repository. Each set of data is described in terms of the location, time of year, time of day, weather, sensors used in the collection, and keywords describing the data. A web query interface is used to select data from the repository (Figure 11 and Figure 12). This interface will shortly be made available outside the NIST firewall, and researchers are encouraged to take advantage of the data.

A wide range of off-road data has been collected, including desert, woods, grassy areas, bushes, water, tree lines, obstacles (rocks, trees, ditches, etc.) and undulations and slopes of various sorts (Figure 7). On-road data includes dirt and gravel roads as well as paved roads, road markings, road signs, and features along the sides of the road (Figure 13). Some of the data includes pedestrians, other vehicles, and special situations such as roadwork, human gestures for guiding the vehicle, and images of a calibration target.

Ground truth data has been acquired for some of the data from an aerial survey of the NIST grounds and surrounding area at a resolution of about 0.3 m per point. The ground truth includes features such as roads, road signs and markings, telephone poles, buildings, trees, fences, ponds, etc. This data provides both a way of evaluating the sensory data and a resource for testing recognition algorithms and using a priori information in sensory processing.

Other information available with the data includes the relative positions and orientations of the sensors, their calibration parameters, the time at which each sample was collected, and the position and orientation of the vehicle at that time. This makes it easy to register the sensors with each other and with the location of the vehicle in the world.


Data Query for Autonomous Mobility

Sensors: Any Location: Any

Vehicle: ○ SUV ○ NIST HMMWV ● Any ○ Other

Sensors of interest:

□ GDRS Ladar □ Color Camera □ Color Stereo □ FLIR Stereo □ Radar □ Acoustic □ Pmgl □ Sick Ladar

□ Other Enter other sensor:

Weather: Any Time of Day: Any

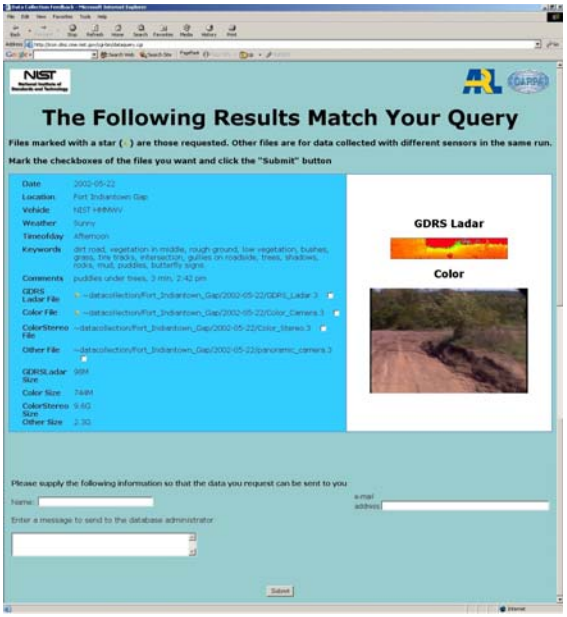
Keywords: (significant features of the data that you want to appear, such as terrain types, ground cover, etc.):

Figure 11 *Web interface for querying sets of data.*

7. DISCUSSION AND CONCLUSIONS

Given a dataset captured in the manner described above, we can borrow the evaluation procedure from the FERET program[9] to quantitatively evaluate the performance of sensor-processing algorithms such as segmentation, classification, and recognition algorithms. These algorithms produce labeled regions in an image. The regions can be projected into the a priori data and assigned labels from the ground truth. It then becomes a simple matter to determine the percentage of false positive and false negative labels of each algorithm and the correctness of the detected positions and shapes of the objects.

The ground truth data are also an excellent resource for verifying the accuracy of a ladar sensor by taking samples from locations that contain surfaces or objects of known sizes, distances, and orientations. The response of the algorithm is then compared with the ground truth position, which is extracted from the database of prior knowledge based on the known position of the sensor and its field of view. Obviously, all measurements are limited by the accuracy of the a priori data and the accuracy with which the position and orientation of the sensor can be established with respect to the a priori data. For the NIST grounds, we have a priori data that are accurate to within a few centimeters, and are working on algorithms to register sensor data with the ground truth to similar accuracy[21]. A sample-by-sample measurement can be made, giving the range resolution and field of view of the sensor. Alternatively, feature-based measurements can be made, giving the accuracy with which the sensor can capture surfaces of different shapes and slopes. More detailed studies, such as trying to determine which part of the field of view of a single sample (e.g., laser beam) gives rise to the measured response, can also be made, but


The Following Results Match Your Query

Files marked with a star ( ) are those requested. Other files are for data collected with different sensors in the same run.

Mark the checkboxes of the files you want and click the "Submit" button

Date	2002-05-22
Location	Fort Indiantown Gap
Vehicle	NIST HMMWV
Weather	Sunny
Timeofday	Afternoon
Keywords	dirt road, vegetation in middle, rough ground, low vegetation, bushes, grass, tire tracks, intersection, gullies on roadside, trees, shadows, rocks, mud, puddles, butterfly signs
Comments	puddles under trees, 3 min, 2:42 pm
GDRS Ladar File	~datacollection/Fort_Indiantown_Gap/2002-05-22/GDRS_Ladar 3 □
Color File	~datacollection/Fort_Indiantown_Gap/2002-05-22/Color_Camera 3 □
ColorStereo File	~datacollection/Fort_Indiantown_Gap/2002-05-22/Color_Stereo 3 □
Other File	~datacollection/Fort_Indiantown_Gap/2002-05-22/panoramic_camera 3 □
GDRSLadar Size	98M
Color Size	744M
ColorStereo Size	9.6G
Other Size	2.3G

Please supply the following information so that the data you request can be sent to you

Name:

e-mail address:

Enter a message to send to the database administrator

Figure 12 *Partial results of query.*

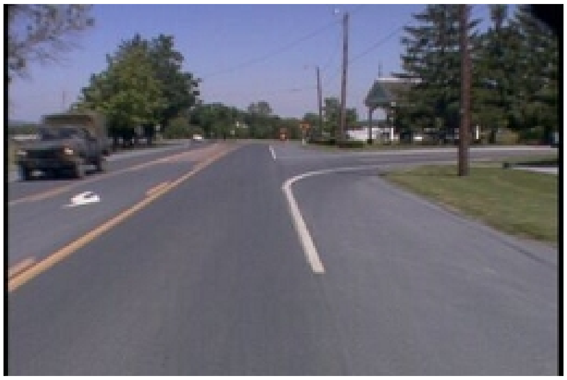


Figure 13 *Sample image from an on-road data set.*

methods customized to the sensor are more reliable.

We have developed a reliable methodology for establishing a large database of ground truth for evaluating sensors and sensor-processing algorithms. The database is available to the public with the hope that researchers and engineers will use it to verify and evaluate sensors and algorithms for effectiveness, efficiency, reliability, and robustness. This will enable algorithms to be developed using realistically difficult sensory data, allow quantitative comparisons of algorithms by using the same data, and spur technology transfer by providing industry with metrics for comparing algorithm performance. It will also help with sensor development by highlighting areas of strength and weakness of current sensors.

Acknowledgements

We are grateful for the support of the Army Research Laboratory and the Defense Advanced Research Projects Agency MARS program. We thank Alan Lytle for his help with the ArcSecond measurement system.

REFERENCES

[1] Courtney, P. Benchmarking and Performance Evaluation . http://www-prima.inrialpes.fr/ECVNet/benchmarking.html . 3-25-1998.

[2] Faber, A. Quality Characteristics of Pattern Recognition Algorithms. http://www.dagm.de/DAGM/ag/wg.html . 5-6-1998.

[3] Lucas, S. IAPR TC-5 Benchmarking and Software. http://algoval.essex.ac.uk/tc5/Introduction.html . 2003.

[4] Bowyer K., Kranenburg, C., and Dougherty, S. Edge Detector Evaluation Using Empirical ROC Curves. Proceedings of the IEEE COmputer Society Conference on Computer Vision and Pattern Recognition, 354-359. 1999. Los Alamitos, CA, IEEE.

[5] Nguyen, T. B. and Zhou, D., "Contextual and Non-Contextual Performance Evaluation of Edge Detectors," *Pattern Recognition Letters*, vol. 21 pp. 805-816, 2000.

[6] Matthies, L., Litwin, T., Owens, K., Murphy, K, Coombs, D., Gilsinn, J., Hong, T., Legowik, S., Nashman, M., and Yoshimi, B. Performance Evaluation of UGV Obstacle Detection with CCD/FLIR Stereo Vision and LADAR. IEEE Workshop on Perception for Mobile Agents. June,1998. Santa Clara, CA.

[7] Shufelt, J. A., "Performance Evaluation and Analysis of Monocular Building Extraction From Aerial Imagery," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 311-326, 1999.

[8] Wiedemannm C., Heipke, C., Mayer, M., and Jamet, O., "Empirical Evaluation of Automatically Extracted Road Axes," *Empirical Evaluation Techniques in Computer Vision* 1998, pp. 172-187.

[9] Phillips, P. J., Moon, H., Rizvu, S. A., and Rauss, P., "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, 2000.

[10] Cho, K., Meer, P., and Cabrera, J., "Performance Assessment Through Bootstrap," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1185-1198, 1997.

[11] Courtney, P., Thacker, N., and Clark, A. F., "Algorithmic Modelling for Performance Evaluation," *Machine Vision and Applications*, vol. 9, no. 5-6, pp. 219-228, 1997.

[12] Kiryati, N., Kälviäinen, H., and Alaoutinen, S., "Randomized or Probabilistic Hough Transform: Unified Performance Evaluation," *Pattern Recognition Letters*, vol. 21, no. 13-14, pp. 1157-1164, 2000.

[13] Haralick, R. Propagating Covariance In Computer Vision. Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms. 1996. Cambridge, England.

[14] Pissaloux, E. E., "Toward an image segmentation benchmark for evaluation of vision systems," *Journal of Electronic Imaging*, vol. 10, no. 1, pp. 203-212, 2001.

[15] Min, J., Powell, M. W., and Bowyer K. Automated performance evaluation of range image segmentation. Fifth IEEE Workshop on Applications of Computer Vision, 163-168. 2000. Palm Springs, CA, IEEE.

[16] Coutre, S. C, Evens, M. W., and Armato II, S. G. Performance Evaluation of Image Registration. Proceedings of the 22nd Annual EMBS International Conference, 3140-3143. July,2000. Chicago, IL, IEEE.

[17] Shin, M. C., Goldgof, D., and Bowyer, K. W. Objective Comparison Methodology of Edge Detection Algorithms Using a Structure From Motion Task. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. June,1998. Santa Barbara, CA, IEEE.

[18] Moon, H., Chellappa, R., and Rosenfeld, A., "Performance Analysis of a simple Vehicle Detection Algorithm," *Image and Vision Computing*, vol. 20, no. 1, pp. 1-13, 2002.

[19] Scott, H. and Szabo, S. Evaluating the Performance of a Vehicle Pose Measurement System. Proceedings of the Performance Metrics for Intelligent Systems (PerMIS) Workshop. Aug.,2002. Gaithersburg, MD.

[20] Bouguet, J.-Y. Camera Calibration Toolbox for Matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/ . 10-2-2002.

[21] Madhavan, R. and Messina, E. Iterative Registration of 3D LADAR Data for Autonomous Navigation. Proceedings of the IEEE Intelligent Vehicles Symposium. 2003. Columbus, OH, IEEE.